TREE SHAKING AND JELLY MAKING: GROWING UP WITH RETROVIRUSES

to be given April 4, 1985, as Faculty Research Lecture

_____, colleagues, friends:

One thing I have learned from the experience of preparing this lecture is the impact of a mystifying title.

We shall come a little later to the origins of the first half of the title, but I want to speak initially about the idea of growing up with and living within a scholarly field. One of the novels that crossed my nightstand in the past year---Disturbances in the Field by Lynn Sharon Schwartz---uses its characters to address explicitly the problems of individual initiative and emotional response within a larger and more powerful network of social behavior. There are similar tensions in our lives as scientists, between the field in which we work---a social and intellectual maelstrom with a life of its own---and our individual aspirations to send out on occasion some positive signal. This dichotomy will be apparent in my talk. I will attempt first to outline broad currents in the study of retroviruses and then give you a partial account of one route I have taken within this network of people and ideas. As you will see, Disturbances in the Field, the title of Lynn Schwartz' novel, could have worked quite nicely for me.

It will be helpful later if a spend a moment now to tell you how I got involved in this field. Two factors were decisive. One was night school at the National Institutes of Health, where medically-

obstacle to the acceptance of the provirus hypothesis. Second, Steve Martin had isolated a temperature-sensitive mutant of Rous sarcoma virus, a mutant that clearly distinguished between functions required for multiplication of retroviruses and a function required for its neoplastic effects. It was apparent to many that a genetic and molecular approach to RNA tumor viruses was now possible. Money flowed, and people too.

(lights down, slide on) During the subsequent decade and a half, two questions have dominated the thoughts of most of us working with retroviruses: How do they reproduce? and How do they cause tumors? As I hope my narrative today will make clear, the answers are intertwined, and they have raised further questions that were inconceivable when we began.

(next slide: exog. retros) Most retroviruses are external agents that invade animal cells. In this sense they resemble viruses of many other kinds that infect cells in all known biological kingdoms. Invasion is followed by multiplication of the virus, but with important differences that distinguish retroviruses from most other viruses. Retroviruses rarely kill their host cells; instead viral genes (shown in red throughout most of the slides) are perpetuated in the infected cells, and the cells may appear either unchanged or given a growth advantage as a result of neoplastic transformation, which also affects the appearance of the cells. For over half a century, tumor virology hobbled along, working mainly with a few exogenous retroviruses of birds and mice, principally those viruses that will recur frequently in our discussion today: the leukosis and Rous sarcoma viruses of chickens and the mammary tumor and leukemia viruses

of mice. But we now know retroviruses to be widely dispersed in nature (first colored tree), found certainly in fish and snakes (and perhaps in non-vertebrates), as well as in several birds, rodents, ungulates, and primates. Among the most recent and best known exogenous retroviruses are the apparent causative agents of two human diseases, adult T cell leukemia and AIDS.

(next slide: info diagram) The biochemical principle that unites all retroviruses was at one time considered an eccentric violation of the usual flow of genetic information from DNA to RNA to protein. Retroviruses encode enzymes, among them reverse transcriptase, that convert their genes, carried about in an RNA form, to a more stable DNA form, the provirus ----a form that can be used by host cell machinery normally entrusted with the job of duplicating and expressing the cell's own genetic legacy. In the past few years, retroviruses have come to seem much less odd. Several other kinds of viruses and genetic elements have been shown or strongly suspected of transferring their information through an RNA form and back to DNA, as a means of duplicating the information and moving it within or between cells. (next fruiting tree: retro elements) Among these are viruses that carry their genes in a DNA form but convert them transiently to RNA during virus growth: the several hepatitis B viruses of birds, rodents, and man and the cauliflower mosaic virus. Other "retro elements" are components of cell chromosomes that have been generated by intracellular events involving reverse transcription. These include a number of mobile genetic elements of yeast, flies, and vertebrates to which I will return; several kinds of repeated sequences; and apparently worthless copies of genes.

The connections that I will emphasize today between retroviruses and certain mobile components of chromosomes were presaged almost twenty years ago by evidence that retroviral genes could be transmitted through the germ line, (next slide: genetic transmission) in a form we now recognize---according to rules I'll discuss later--- to be endogenous proviruses. Such proviruses are components of normal chromosomes, passed on to future progeny, and presumed to be the consequences of infection of germinal cells in the recent or distant past. In somatic cells, endogenous proviruses may be completely silent or they may be expressed. Some can generate fully infectious viruses, and the endogenous viruses may even be tumorigenic. Most of these proviruses, however, appear to be functionally incompetent. Though first studied in mice and chickens, inherited, provirus-like elements are now known to be abundant in virtually all eukaryotes (next fruited tree), present in yeast, flies, and all vertebrates including man. There are, at a minimum, a few thousand such elements in our chromosomes, approaching one percent of our total DNA. What are they doing there? It is anyone's guess. The pundits offer many possibilities: the elements may hasten evolutionary change and hence be good for us; they may cause considerable damage but be tolerated if they do not; or their existence may be largely unrelated to effects upon their hosts, and instead reflect only the elements' selfish interest in survival.

(next slide: expression) Whether exogenous or endogenous, retroviruses exploit mechanisms provided by their host cells and contribute their own idiosyncratic mechanisms to produce viral RNA, protein, and ultimately infectious particles from proviral DNA.

Consequently, (next fruited tree) retroviruses have been important reagents in efforts to learn how eukaryotic cells regulate their genes---through synthesis of RNA, with promoters, enhancers, hormonal regulators, and splicing; through synthesis and modification of proteins; and through coordinated assembly and release of virus particles competent to infect a new host cell efficiently.

A closer look at retroviruses reveals an extraordinary potential for complex genetic interaction with each other and with their hosts. (next slide) Retroviruses carry two RNA copies of their genes---hence, like us, but unlike other viruses, they are diploid. When two distinguishable retroviruses infect the same cells a number of consequences are possible: heterozygotes, with two different sets of genes; recombinants, with new combinations of genes; or pseudotypes, with the proteins of one virus wrapping up the genes of another. As we shall explore in greater detail later, the interaction between viral and host genes is even more remarkable, with proviral DNA covalently joined to the host chromosome. This may lead to mutation of host genes and occasionally to incorporation of host genes into progeny viruses, a process called transduction.(next fruiting tree) These phenomena (and others) provide the principles for a new phase of retrovirology, in which retroviruses are modified by design for use as genetic vectors. Within the next year or two, we will likely see the first attempts to cure certain inherited human diseases with such vectors.

(next slide: transduction) The most potent of the oncogenic retroviruses are naturally-occuring genetic vectors, having captured host genes during passage of wild type retroviruses through cells.

The transduced genes are responsible for the neoplastic activities of these viruses and are known as viral oncogenes or v-onc's; their cellular progenitors are called c-onc's or proto-oncogenes. Following the elucidation of these principles with the src gene of Rous sarcoma virus a decade ago, (next fruited tree) src has been joined by about twenty other genes, some closely related and some not, some first encountered in other highly oncogenic viruses of birds, some in viruses of mice, rats, and cats, one in a virus of monkeys. Each of these viral oncogenes is derived from a cellular proto-oncogene, genes that are carefully conserved in evolution---and hence found in all vertebrates and sometimes in insects, sponges, yeast, and other organisms. These cellular progenitors of viral oncogenes are in one sense misnamed: (next slide) under normal conditions, proto-oncogenes appear to supply functions essential for the growth, development, and function of cells. They provide extracellular factors that signal cells to divide; cell membrane proteins that serve as receptors for such growth factors and possess protein kinase activity specific for tyrosine residues in target proteins; other kinases, some also specific for tyrosine, but without known physiological function; GTPases that are presumptive modulators of external signals for growth; and nuclear proteins that may regulate the expression of other genes. (next slide) On the other hand, in many cancers of man and animals, proto-oncogenes fulfill the promise of their name, having been converted to apparently active oncogenes by a variety of mutational mechanisms, including gene amplification, chromosomal translocations, nearby insertion of mobile elements, rearrangements that truncate genes, or simple point mutations. Naturally, these

observations have been the seeds for additional trees (next slide), one of which is illustrated by the kinase family of oncogenes and proto-oncogenes.

(composite tree slide) I have tried to convince you that the orchard of retrovirology has flowered and fruited marvellously over the past fifteen years, nourished in some collective sense by those of us who work in it. As we turn to consider individual contributions to the discipline, I find myself shifting metaphors slightly, thinking more about harvesting than nurturing, but wondering what determines how each of us fares in this garden. To what extent does a good crop signify character, luck, insight, or a transient stage of life or career? (slide off; lights)

It was a coincidence of vague thoughts of this sort with a small item in last November's newpaper that provided the mysterious title of today's talk. Following the defeat of the Democrats in last year's election, some of Jesse Jackson's supporters were asked what he planned to do with the public platform he had built during the campaign. "We want him to be a tree-shaker, not a jelly-maker," one of them replied. The reporter went on to explain that Jackson was expected to speak and act in ways that knocked issues out of the trees so they could be seen and acted upon; the pitting, boiling, sugaring, and bottling (read litigation, legislation, enforcement) would be left to others less imaginative.

These are concerns that I suspect many of us have about our work. We all seek those startling results that shift paradigms and found new disciplines. But is it possble to make a meaningful distinction between tree-shaking and jelly-making in daily laboratory

life? Is it possible to start each day with a revolutionary plan? Or are we just as likely to meet with thunderous success by hewing closely to Pasteur's dictum, "Chance and the prepared mind"? As we grow with a discipline, we become attached to our own findings and our own reagents---cell lines, mutants, techniques ---often to the point at which it is legitible to wonder, "Would I be trying to solve the problem at the heart of my research efforts in some better way if I now entered it anew and attacked it without the baggage of the past? "

I don't see simple answers to these questions. I have had the good luck to be associated some years ago with one episode of what now seems to be a candidate for tree-shaking: the discovery that retroviral oncogenes are derived from normal cellular genes. Yet even that discovery was not instanteously perceived as tree-shaking by us---and certainly not by our audience---until a lot of jelly-making, and even some additional tree-shaking, had occurred in a large number of labs.

I would like to explore these considerations further by telling a story that begins with a simple and potentially specialized (i.e. jelly-making) question about how retroviruses reproduce. The tale leads first to a description of proviral DNA that links retroviruses with other kinds of mobile genetic elements; it then provides the first evidence that cellular proto-oncogenes can become true oncogenes without conversion to viral oncogenes; and it ends with the discovery of an novel oncogene implicated in mammary carcinomas of mice.

By 1975, the provirus hypothesis was firmly established. (lights down: slide) Work here and elsewhere had shown that viral DNA is made in the cytoplasm of retrovirus-infected cells by reverse transcriptase

and later integrataed covalently into host chromosomes, where it then serves as a template for synthesis of viral RNA. Ram Guntaka had recently discovered here that retroviral DNA is present in a circular as well as a double stranded linear form early after infection. The model shown was proposed at that point for early events in the virus life cycle, showing linear DNA, approximately the size of its viral RNA template, as precursor to circular DNA, circular DNA integrating into host DNA, and proviral DNA being transcribed into viral RNA. Attention was focused upon the organization of viral DNA, the manner in which it was synthesized and integrated, and the positions in host chromosomes at which integration occurred.

This scheme was correct in general. But thinking about some crucial details was troubled by the following dilemma. (next slide) Several laboratories had examined the initial steps in reverse transcription by working with the viral enzyme in vitro. The primer for DNA synthesis was found to be a host tRNA bound to viral RNA near its 5' end. Synthesis of DNA could proceed for only a very short way, about 1% of the RNA template, before something unusual had to occur. Because viral RNA was shown to be terminally redundant---that is, to have exactly the same short sequence, called R, at each end---and because reverse transcriptase is associated with another enzymatic activity that can remove RNA from an RNA:DNA hybrid, it was proposed (and later proven) that the DNA copy of R (called R') is exposed to form a hybrid with the other (3') end of viral RNA. This allows the nascent DNA to prime synthesis along the rest of viral RNA. While it solves the problem of extending the DNA strand, this manuever creates another problem: namely, a copy of R has apparently been sacrificed.

(slide) At first glance this would be disasterous for the virus life cycle: copies of R must be present at both ends of proviral DNA to provide an appropriate template for synthesis of viral RNA. A copy of R must be regained, but when and how? It is sobering to reveal to you the answers that I proposed at the time to my colleagues in public places. We then called the R sequence X, and the schemes included (slide) integration at a host site that contained a second copy of X, (slide) staggered cutting of circular DNA to duplicate X, and (slide) tandemly integrated proviruses, each providing a single copy of X. Each of these seems reasonably clever and would solve the problem during the integration step; but all are wrong and essentially less interesting than the real answer.

Finding the answer was technically demanding---remember these not-so-long ago days preceded molecular cloning of retroviral DNA---and it represented one of the first applications of Southern's DNA transfer procedure to eukaryotic sequences present in very few copies per cell.(slide) The successful strategies made use of molecular probes specific for the region of viral RNA positioned between the tRNA primer and R---a sequence known hereafter as U5---and for a region situated near the other copy of R (and known as U3). We were also helped by recognition signals for certain restriction enzymes in these regions (e.g. an Eco RI site in U3).

The picture of unintegrated linear DNA that emerged from this work (performed by Peter Shank, Steve Hughes, and HJ Kung) was extraordinary and unexpected (slide): both U3 and U5 (and, by implication, a copy of R) were present at both ends of linear DNA. This meant that the copy of R was restored during DNA synthesis rather

than during integration; that sequences present only once in viral RNA (i.e. U3 and U5) were present twice in DNA; that the arrangement of these sequences created terminal repeats in the DNA that were substantially longer than the terminal repeat in the RNA. Furthermore, creation of these long repeats required a second acrobatic event during DNA synthesis, one resembling the first jump between templates described earlier. (This event, the second jump, is now understood but a distraction from the story I want to tell.)

(next slide) Linear DNA was then shown to be converted to two forms of circular DNA, a form with one copy of the long terminal repeat (or LTR) and one with two copies that was recently proven to be the integrating species. Integrated proviral DNA itself was found to resemble the linear DNA, save for the fact that it was joined to host DNA at both ends, with each provirus linked to a different part of the host genome. According to this strategy, proviral DNA was beautifully equipped to be a template for production of viral RNA--- the R sequence was present near both ends, and there was additional sequence that might encode regulatory signals, such as transcriptional promoters, and would reside upstream of the appropriate initiation site. The results also implied that integration of viral DNA was quite precise, always joining the same region of viral DNA to cellular DNA, but random or quasi-random with respect to the cellular site at which integration occurs.

With later application of molecular cloning and nucleotide sequencing techniques, each of these deductions was confirmed and additional symmetries were perceived. (next slide) First, the LTRs were found to have internal symmetries, with their terminal

nucleotides arranged as inversions of the same sequence. Second, the site in viral DNA at which linkage to host DNA occurs is precise to the nucleotide, with exactly two base pairs lost from each LTR. (This loss is tolerated because it involves a region that is not transcribed into RNA.) Third, though the host integration sites are always different for each provirus, a few nucleotides at the insertion site are precisely duplicated to flank the provirus.

Finally, (next slide: LTR anatomy) both structural and functional studies show that the LTR is truly the hub of the proviral universe---defined by initiation sites for synthesis of both DNA strands, containing the sites for integration, and equipped with signals to start RNA at the upstream end of a provirus and to end it at the downstream end.

In mid-1978, just after LTRs had come into view, we moved to London to spend a sabbatical at the Imperial Cancer Research Fund and in a Georgian row house in Islington (slide). Having plenty of time to ruminate, I began to think seriously about the possibility that integration of retroviral DNA would, in effect, mutate the host chromosome. (next slide: structural homologies) This idea had actually been self-evident from the time proviruses were first proposed to integrate into host DNA, but it took added force from description of the LTRs. It was now apparent that the proviruses of retroviruses---here represented by the two we had studied, Rous sarcoma and mouse mammary tumor virus---were affiliated, at least on structural grounds, with mobile genetic elements of bacteria, yeast, and Drosophila. (The latter, as you have already heard, are also functionally related to retroviral proviruses, since they pass through

RNA intermediates.) In general, when these elements move (or transpose) they enter host chromosomes promiscuously; (next slide: insertion muts) hence they are prone to make functionally significant mutations, either by disrupting and inactivating a normal gene or by augmenting its expression from a nearby insertion site.

I decided to show directly that retroviruses can act as mutagens by interrupting a pre-existing gene in a cellular chromosome. I chose to use as a genetic target for an invading provirus not a cellular gene, but instead (next slide) the single Rous sarcoma provirus whose src gene and protein product had been responsible for neoplastic transformation of a normal rat cell, producing the line called B31. Superinfection of B31 cells with another retrovirus---a murine leukemia virus that lacks an oncogene and does not transform cells in culture---should revert the behavior of B31 cells to normal in those rare cells, roughly one in a million, whose Rous provirus is interrupted by a newly acquired murine leukemia provirus. Moreover, if the latter provirus were then excised, the cells might return to the transformed state. (next slide: ins1,2) The expected mutants were indeed rare, but Suzanne Ortiz and I were able to isolate two of them, and they exhibited the predicted properties. In both cases, an MLV provirus was stuck in the middle of the RSV provirus---interestingly, not in the coding region for src but in the middle of a region normally removed by splicing to produce src messenger RNA. As a result, little or no src mRNA or protein was found in the mutant cells and the cells lost their transformed phenotype. In one of the two lines, most of the MLV provirus was eliminated at low frequency, by recombination between the LTRs. This rare event was readily

detectable, however, because it restored expression of the src gene and the transformed phenotype. In short, retroviruses were proven competent to make insertion mutations that inactivate genes, at about the frequency expected for a randomly integrating element. (This conclusion has now been corroborated by other laboratories, in both germinal and somatic cells, using true host genes as targets.)

But what about insertion mutations that activate cellular genes? About the time of my sabbatical, Nancy Quintrell performed an experiment that I now like to describe as heuristic. (slide of NQ's expt) In surveys of viral DNA and RNA in Rous sarcoma virus transformed mammalian cells, she happened upon one cell line in which the provirus appeared to be promoting transcription of adjacent cellular DNA, since a novel messenger RNA was shown to contain host sequences linked to sequences---R and U5---from the viral LTR. No matter that the conclusion shown here is inaccurate in certain details, or that the phenomenon had no apparent consequences for this cell. Its effect was instead upon our thinking about another situation: (slide) how does a retrovirus that lacks an oncogene manage to induce a tumor?

I indicated earlier that not all oncogenic retroviruses carry viral oncogenes---genes derived from cellular proto-oncogenes. Those that don't, tend to produce tumors sluggishly and are incompetent to transform cultured cells. Since these viruses are vigorously infectious, it seemed plausible to ask whether their LTRs might drive inappropriate expression of some proto-oncogene that happened to reside near (presumably downstream from) an inserted provirus. Even if this event occurred extremely rarely, it should be detectable when the affected cell grew clonally to form a tumor.

Greg Payne, then a graduate student in our group, considered this possibility in his work with the avian leukosis virus--- a chicken retrovirus devoid of a viral oncogene but still able to produce tumors, most commonly lymphomas in the B cell organ of the chicken, the bursa of Fabricius. (slide: models and LL1) Some tumors did in fact contain species of RNA similar to that Nancy had described, with LTR sequences linked to unidentified host sequences. Moreover, such tumors---for example, tumor LL1--- usually had only a single provirus and the provirus was often aberrant, so that it was not possible to express viral genes. This suggested the viral genes used for virus replication were not important for tumor growth, but whatever sequences resided downstream from the provirus might be.

It was obviously critical to know what those sequences were. Physical mapping with restriction enzymes, both by Greg and in parallel work by Hayward, Astrin, and their colleagues, suggested that the activated host sequences might be the same in different tumors. At this point Hayward caused what I would now call "a disturbance in the field." (I still lose sleep over this one, though I assume Hayward does not.) He used available probes for several retroviral oncogenes, and found (slide) that the vast majority of bursal lymphomas contained ALV proviruses adjacent to a now efficiently transcribed cellular myc gene, the proto-oncogene previously discovered by Mike and Diana Sheiness when they traced the cellular origins of the v-myc oncogene of a highly oncogenic avian retroviruses.

Here was a tree-shaker: for the first time, cellular proto-oncogenes appeared to deserve their perjorative name after all: they

could be activated in their original settings and thereby contribute to neoplasia. A few of you have heard me speak previously about the musicality of discovery. For example, the finding that viral oncogenes are derived from cellular genes seems Wagnerian in texture: the piece draws much of its strength from repeated restatement of the theme. Hayward's experiment was Mozartian in character: a single, transporting phrase, one that seemed effortless but suggested infinite possibilities. The physical execution of his experiment was trivial; we were able to confirm the results within days of his phone call. Why didn't we do it earlier? It will cost you a beer to find out.

When we regained our composure, we realized that our collection of tumors differed in an interesting way from Hayward's: (next slide: diagram of myc) all of his tumors contained proviruses positioned so that an LTR could provide a promoter for c-myc. However, several of our tumors had proviruses arranged differently within the c-myc locus. In such cases, with the provirus upstream from c-myc but in the opposite transcriptional orientation or with the provirus downstream from c-myc, another property of the provirus must be responsible for augmented expression. This was one of the first pieces of evidence that proviruses contained transcriptional enhancers, sequences that improve the performance of adjacent promoters over relatively large distances and in a fashion independent of orientation.

Further study of the mutated c-myc loci by Greg and Dave Westaway showed that the insertion of ALV DNA was not the sole lesion: usually the proviruses were damaged by deletions and sometimes the coding sequence of c-myc was also altered. These secondary mutations are consistent with clinical opinion that the development of a cancer

requires multiple rare events, reinforcing our confidence that tumor induction by viruses without oncogenes is a credible model for human cancers.

This episode with bursal lymphomas has had a dramatic effect upon many aspects of retrovirology and oncology. (slide: rearrangements) The most far-ranging influence of insertion mutations in the c-myc locus has been to serve as harbingers of other kinds of rearrangements of DNA, alterations of chromosomal architecture that affect other proto-oncogenes as well as c-myc, in both human and experimental tumors. This slide illustrates some well known examples --- chromosomal translocations in which c-myc genes in Burkitt lymphomas are fused to immunoglobulin genes on other chromosomes, and localized amplifications of cellular DNA that augment the number of copies of c-myc genes and the abundance of the gene products.

Proviral insertions in c-myc also stimulated our thinking about the mechanism by which retroviruses capture host proto-oncogenes and convert them to viral oncogenes. (slide) As shown in this widely-accepted scheme, whose principal architect was Ron Swanstrom, transduction is proposed to begin with a proviral insertion mutation upstream from a proto-oncogene and in the same transcriptional orientation. I like to think that this event initiates uncontrolled cell growth, expanding the number of cells in which the subsequent rare events might occur. A deletion mutation then removes the downstream part of the provirus, making a hybrid transcriptional unit. After splicing, messenger RNA from such units would have one end that resembles normal retroviral RNA and includes signals for packaging into virus particles. The other end contains some of the exons from

the cellular proto-oncogene. If another normal provirus is in the same cell, heterozygote genomes can form, and recombination is likely to follow. The final product is an RNA molecule with all the features required for replication as a retroviral genome, provided that viral proteins are supplied by an accompanying helper virus. This is the case for the vast majority of retroviruses bearing oncogenes. Several factors may contribute to the physiological differences between an oncogenic v-onc and a benign c-onc: sustained and efficient expression from viral LTRs, truncations of the coding regions, and more subtle nucleotide sequence changes that occur during virus multiplication.

Let's pause briefly here to review where we have been. You have seen how proviruses were found to resemble transposable elements, and I have shown you examples of three sorts of proviral insertion mutations: (slide: IM1) mutations that inactivate genes, (slide: IM2)mutations that activate genes by promoter insertion and could represent the first step in the transduction of oncogenes, and (slide: IM3) mutations that activate genes by an enhancer mechanism. One of the rewards of mutations should be the identification and isolation of new genes, something I have yet to describe. This brings me to the final stage of today's story.

(slide of C3H mouse) Our group has been interested since the early 1970's in the mammary carcinomas caused by the mouse mammary tumor virus; (slide: summary) MMTV is a retrovirus commonly transmitted through milk to nursing mice. Like avian leukosis virus, it lacks a viral oncogene, and induces clonal neoplastic growths that contain new proviruses, suggesting that an integration event may be important in carcinogenesis. Finally, the MMTV LTR harbors an

unusual promoter with glucocorticoid responsiveness, suggesting that an atypical regulatory phenomenon may also be important.

In 1981, Roel Nusse and I set out (slide: bicyclists) to ask whether MMTV proviruses were causing insertion mutations in these tumors (slide: ?) and, if so, to isolate the affected proto-oncogene--- hoping that we would be led to a novel gene, not back to c-myc or some other known proto-oncogene. In the wake of Hayward's results, the latter would only be interesting jelly-making.

Our strategy was similar to one that had been recently used to isolate an eye color gene from Drosophila melanogaster. The mobile genetic element, copia, had caused a mutation in the white locus; by cloning that copy of the copia element, it was possible to gain a foothold in the gene, because a piece of it was adjacent to copia in the molecular clone. (slide) To practice this art, now known as transposon tagging, Roel sought and found a mouse mammary tumor bearing only a single new MMTV provirus. According to the hypothesis, this provirus should be adjacent to an activated proto-oncogene. If tumorigenesis depends upon the rare insertion of viral DNA near this proto-oncogene, other tumors should also have proviruses somewhere in the vicinity. To test these predictions, part of the provirus was cloned from the initial tumor with flanking host DNA still attached; the host sequences were then separated from viral DNA and used as a probe to ask two questions: whether the integration site was interrupted in other tumors as a result of MMTV insertions, and whether the insertions altered expression of any genes in the region.

(slide: int insertions) This diagram summarizes the conclusions drawn from many experiments. About three-fourths of mammary

carcinomas in C3H mice have MMTV proviruses inserted within a 20-30 kb domain, roughly equally distributed on both sides of a gene that we now call int-1. Though silent in normal mammary tissue, the gene is expressed when an MMTV provirus is nearby. You will notice that the proviruses are not positioned to allow the LTR to act as a promoter for int-1; virtually all the proviruses are either upstream from int-1 in the opposite transcriptional orientation or downstream in the same orientation. We do not yet know why this is so, but the result implies that an MMTV enhancer function is fundamental to its oncogenic potential.

What do we know about the int-1 gene? First, it appears, as hoped, to be an entirely novel proto-oncogene, unrelated to any of the known progenitors of retroviral oncogenes. [The only gene to which it bears reasonable comparison is another gene, called int-2, that a group in England found to be activated in about half of another collection of MMTV-induced mammary tumors.] (slide: int gene) Structural examination of the gene--- by Roel back in Amsterdam--- and of cDNA clones copied here from int-1 mRNA shows that the gene is composed of at least four exons and encodes a protein of moderate size, with a very hydrophobic amino terminus and a cysteine-rich carboxyterminus. The biochemical function of the protein is presently obscure, but we are inclined to believe it is important to normal animals, since, like other proto-oncogenes, it is highly conserved in evolution: Roel has found that the product of the human int-1 gene differs in only four of four hundred amino acids from the mouse int-1 protein. Thus far, there is no evidence that int-1 has a role in human breast cancer. (lights on)

This is as far as I can take you on the circuitous path that began with a search for a short nucleotide sequence lost during reverse transcription and has led to a search for the function of a mammary oncogene. Is there a message in all this about how to foretell jelly-making from tree-shaking? I doubt it. Good luck to Reverend Jackson.

I would like to conclude with a few words about the pleasures of scientific life. The first faculty research lecture I remember attending was given by Gordon Tomkins. He began by claiming that science is fundamentally measurement, and that the first scientist was the caveman who wondered how long it took for a rock to fall from his hand to the ground. I was and am enamoured of this idea, but I am also struck by our need to interpret and represent the things we have measured. I imagine Gordon's first scientist finding complete satisfaction only when he tried to tell a colleague what the speed of the stone's fall meant for the organization of his world. The diagrams that have illustrated this lecture have served a similar purpose for me, reducing an awful lot of numbers and radiographic signals to convenient metaphors for the world of viruses and cancers we seek to understand.

When Tom Sawyer learned how to persuade his friends to paint his fence, he "comprehended that Work consists of whatever a body is obliged to do, and that Play consists of whatever a body is not obliged to do." One of the great pleasures in our lives as academic scientists is the priviledge to engage in extended play. I hope it is not too obviously self-serving to say that it is the mark of an enlightened society---and university---to recognize and subsidize such play, be it jelly-making or tree-shaking.

A word about my playmates. (final fruited tree) Science has long ceased to be a solitary endeavor. Among the best things retroviruses have done for me has been to foster collegial relations with an amazingly large and admirable group of people, including faculty, research staff, and trainees of various sorts. The artist failed to provide enough fruit for everyone; what is shown is a partial accounting of those past and those persistant. The one name that must be spoken is that of Suzanne Ortiz; without her help, I would not have a story to tell.

Finally, (slide:final tree) a word of thanks to the committee that invited me to give this talk and to those of you who came to listen. Some months ago, my wife told me that Bessie Smith used to sing "If you don't want my peaches, don't shake my tree." I am honored that you chose to shake mine.